

Chris Howard Office of the CTO, IBM SWG Europe
Cross Industry Big Data



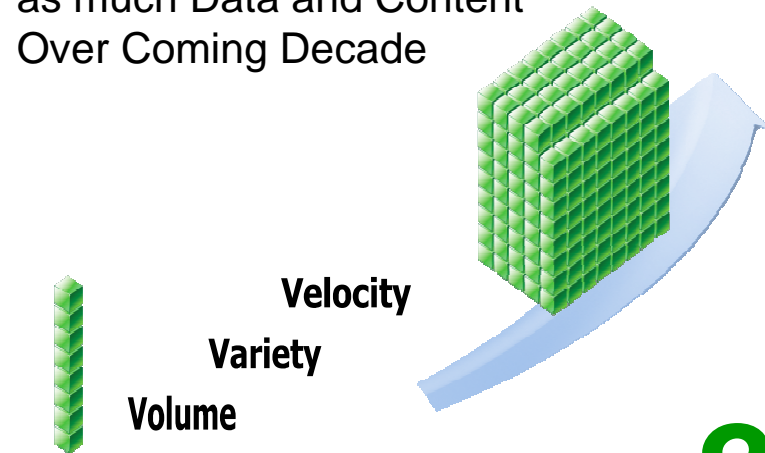
Information is at the Center
of a New Wave of Opportunity...

... And Organizations Need
Deeper Insights

44x

as much Data and Content
Over Coming Decade

2020
35 zettabytes



2009
800,000 petabytes

80%

Of world's data
is unstructured



1 in 3

Business leaders frequently
make decisions based on
information they don't trust, or
don't have

1 in 2

Business leaders say they don't
have access to the information
they need to do their jobs

83%

of CIOs cited "Business
intelligence and analytics" as
part of their visionary plans
to enhance competitiveness

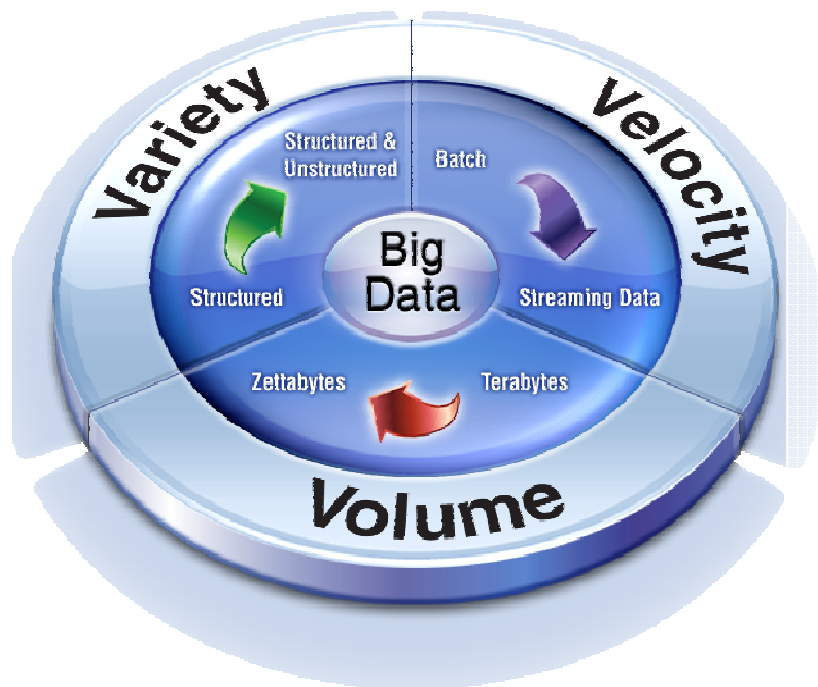
60%

of CEOs need to do a better job
capturing and understanding
information rapidly in order to
make swift business decisions



Big Data Presents Big Opportunities

Extract insight from a high volume, variety and velocity of data in a timely and cost-effective manner



Variety: Manage and benefit from diverse data types and data structures

Velocity: Analyze streaming data and large volumes of persistent data

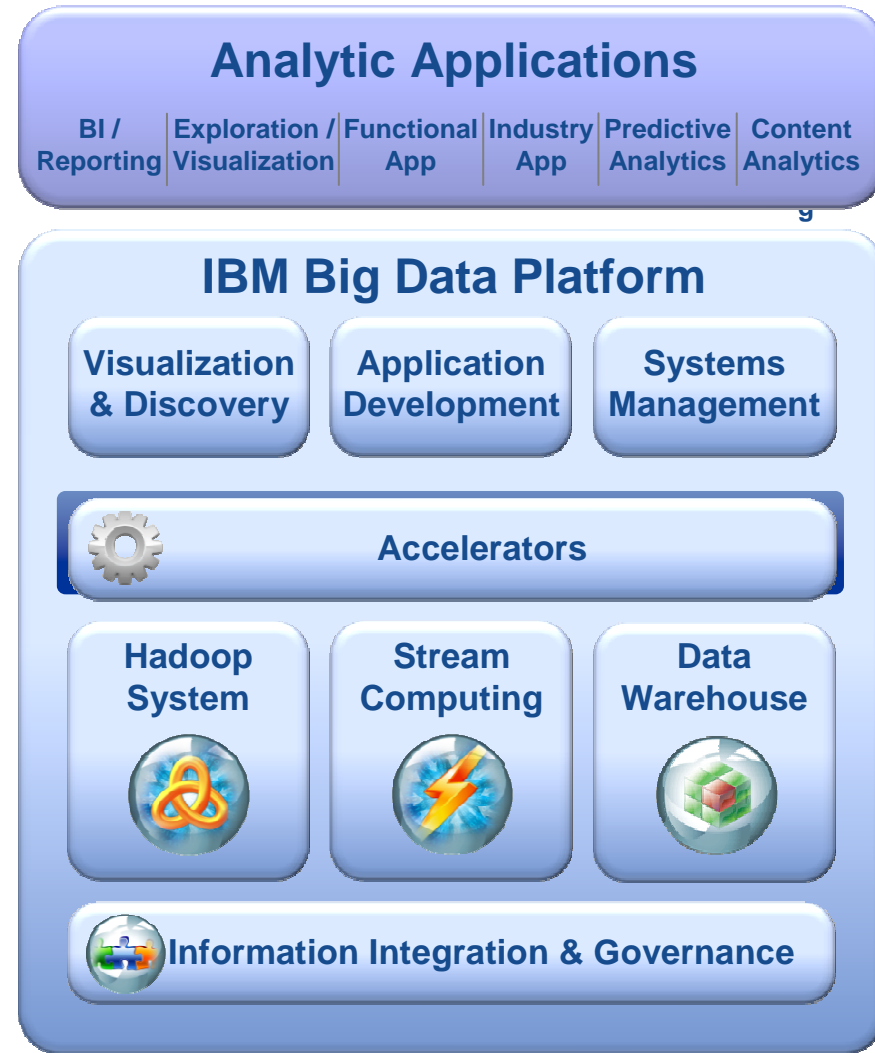
Volume: Scale from terabytes to zettabytes



IBM Big Data Strategy: Move the Analytics Closer to the Data

New analytic applications drive the requirements for a big data platform

- Integrate and manage the full variety, velocity and volume of data
- Apply advanced analytics to information in its native form
- Visualize all available data for ad-hoc analysis
- Development environment for building new analytic applications
- Workload optimization and scheduling
- Security and Governance



Big Data Scenarios Span Many Industries



Multi-channel customer sentiment and experience analysis



Detect life-threatening conditions at hospitals in time to intervene



Predict weather patterns to plan optimal wind turbine usage, and optimize capital expenditure on asset placement



Managing and optimizing real-time multi-modal route planning



Identify criminals and threats from disparate video, audio, and data feeds



Smarter Wind Energy



The Challenge

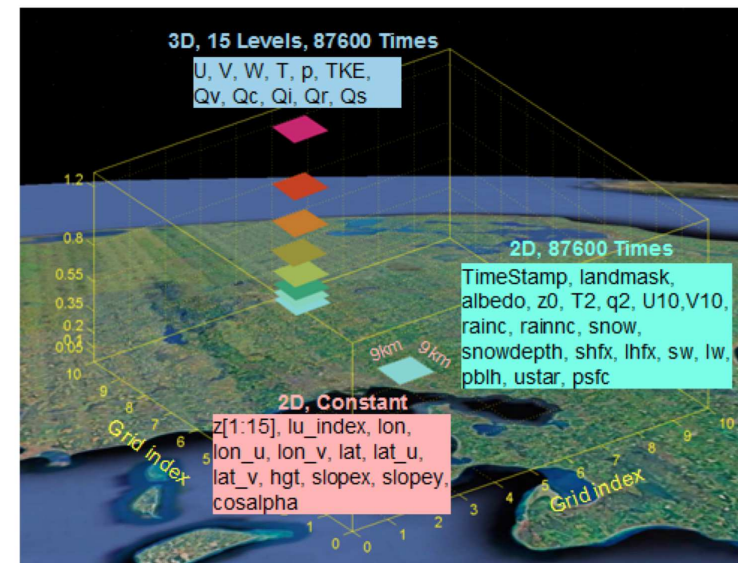
- European target - 20% Renewable energy by 2020
- Analyze large volumes of public and private weather data for alternative energy business
- Leverage large **volume** (2.8+ PB growing to 20 PB) of weather data to optimize placement of turbines – complex discipline
- Combat issues with volatile nature of wind – need to provide customers with confidence that placement site will yield a reliable source of energy – this require complex analytics



Complex Analytics – not just about how hard the wind blows

- Weather is very important to wind turbines!
 - Temperature
 - Snow / ice
 - Yield of the wind turbine over its life ...
- They have their own meteorologists
 - They predict weather on their own
- They run a lot of weather simulations
- They have weather data from the entire world
 - For the last 10 years
 - 35,000 metrological stations worldwide
 - 2 PBytes of data

Data dimensions, 10 years, hourly, any grid resolution



Results and Benefits

- Vestas customers need to build and validate strong business case (need predictability / reliability)
- Modeling time (posing and answering questions) has been reduced from 3 weeks to 15 minutes
- Solution is allowing efficient exploitation of 15,000 core cluster for multiple uses
- Solution provides for significant future scaling
 - Grid resolution
 - Data volumes
- Now investigating optimization of ongoing operations





The Challenge

- Meet the demands of citizens across the city looking to utilize an array of public transportation options
- Optimize routing and scheduling of public transportation to meet passenger demand throughout the day and make effective use of the fleet
- Leverage complex network of sensors, route data and schedules to build a single holistic picture



Routes & maps



Parking capacity



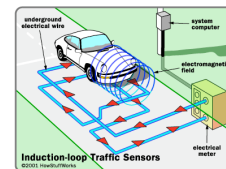
Car



Accessibility



Timetables



SCATS
Induction loop



CCTV



Bike



Complex Analytics

- High-Performance System and Analytics to:
- Assimilate: Subscription to & decoding of SIRI¹ real-time data feeds
- Mediate: De-noising, map matching & trajectory tracking
- and Aggregate: Extraction of key traffic metrics; e.g. speed

¹ SIRI is an XML protocol to allow distributed computers to exchange real-time information about public transport services and vehicles. The protocol is a CEN standard, developed and used by several EU countries

See <http://leicestertravel.info/> or <http://www.tfl.gov.uk/> for examples



Screenshot #1

IBM Transportation Awareness and Optimization System

IBM | IBM Research

- Dashboard
- Bus Delay
- Bus Congestion Count
- Bus Lane Speed
- On-Route Bus Count
- System Monitor
- Search

Map Data: The satellite map shows a city grid with bus routes highlighted in green and red. Yellow bus icons are scattered across the routes. A scale bar at the bottom left indicates 1314 m. The bottom right corner features the Google logo and copyright text: © 2011 Infoterial Ltd & Bluesky.

Charts:

- Average delay (minutes) vs Time:** A line graph showing a constant value of 1.0 from 14:20 to 14:30.
- Number of Vehicles (Log) vs Delay (minutes):** A bar chart showing the distribution of vehicles across different delay ranges. The x-axis ranges from -10 to 10 minutes, and the y-axis is on a log scale from 10 to 100.
- Average speed (km/h) vs Time:** A line graph showing a fluctuating average speed between approximately 15 and 25 km/h.
- Number of Links (Log) vs Delay (minutes):** A bar chart showing the distribution of links across different delay ranges. The x-axis ranges from -10 to 10 minutes, and the y-axis is on a log scale from 10 to 100.



Screenshot #2

IBM Research

Bus stop 1480
Upper Ormond Quay, Ormond House

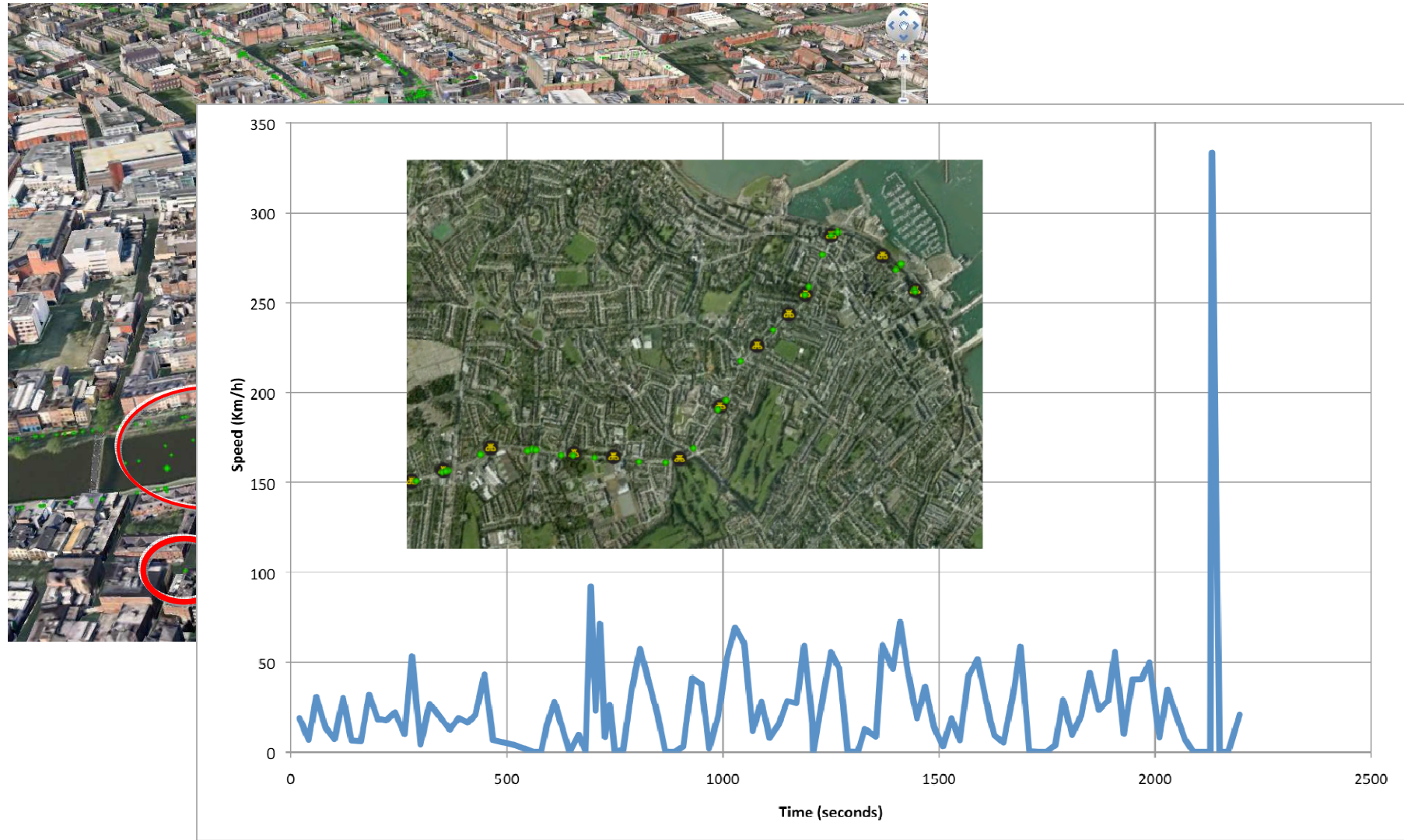
Route	Destination	ETA
37 (var. 001)	Hawkin's Street	7:08 min
39 (var. 001)	Hawkin's Street	0:04 min
51 (var. 001)	Aston Quay	
68 (var. 001)	Crampton Quay	0:03 min
68 (var. 002)	Crampton Quay	17:27 min
68 (var. 003)	Crampton Quay	0:20 min

© 2011 Infoterra Ltd & Bluesky

©2010 Google

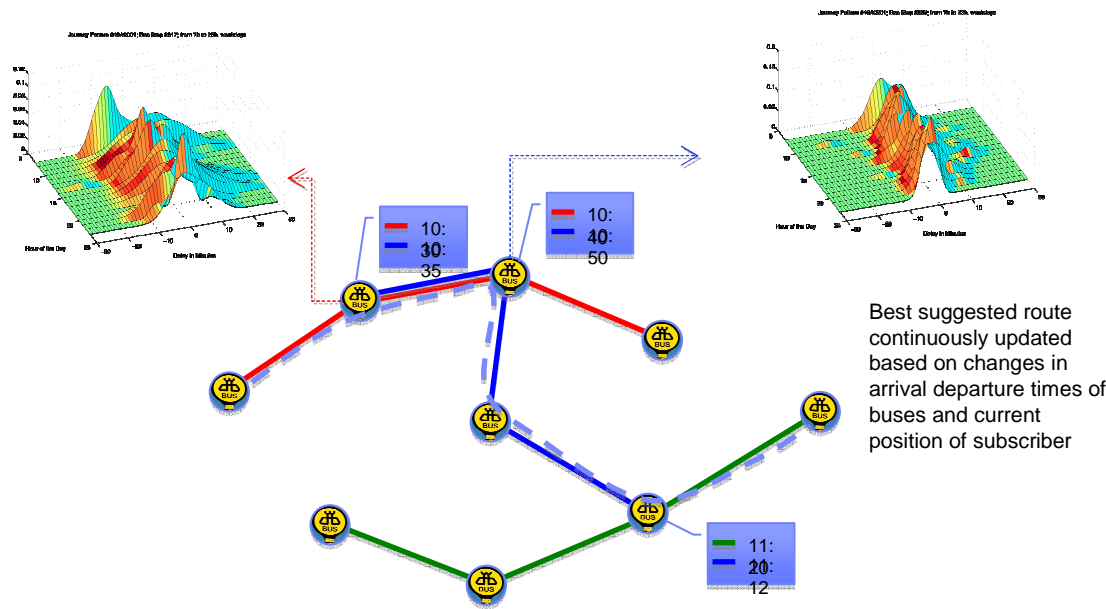


Real-time Sensor Data Challenges



Results and Benefits

- Real-time situational awareness
- Route optimisation – actuating the city
- Real-time multi-modal planning





THINK BIG

